Discrete time McKean-Vlasov control problem: a dynamic programming approach

Huyên Pham, Xiaoli Wei

Discrete time McKean-Vlasov control problem: a dynamic programming approach

Huyên PHAM
Laboratoire de Probabilités et
Modèles Aléatoires, CNRS, UMR 7599
Université Paris Diderot
pham at math.univ-paris-diderot.fr
and CREST-ENSAE

Xiaoli WEI
Laboratoire de Probabilités et
Modèles Aléatoires, CNRS, UMR 7599
Université Paris Diderot
tyswxl at gmail.com

November 29, 2015

Abstract

We consider the stochastic optimal control problem of nonlinear mean-field systems in discrete time. We reformulate the problem into a deterministic control problem with marginal distribution as controlled state variable, and prove that dynamic programming principle holds in its general form. We apply our method for solving explicitly the mean-variance portfolio selection and the multivariate linear-quadratic McKean-Vlasov control problem.

MSC Classification: 60K35, 49L20

Keywords: McKean-Vlasov equation, dynamic programming, calculus of variations.

1 Introduction

The problem studied in this paper concerns the optimal control of nonlinear stochastic dynamical systems in discrete time of McKean-Vlasov type. Such topic is related to the modeling of collective behaviors for a large number of players with mutual interactions, which has led to the theory of mean-field games (MFGs), introduced in [11] and [10].

Since the emergence of MFG theory, the optimal control of mean-field dynamical systems has attracted a lot of interest in the literature, mostly in continuous time. It has been first studied in [1] by functional analysis method with a value function expressed in terms of the Nisio semigroup of operators. More recently, several papers have adopted the stochastic maximum principle for characterizing solutions to the controlled McKean-Vlasov systems in terms of adjoint backward stochastic differential equations (BSDEs), see [2], [6], [7]. We also refer to the paper [16] which focused on the linear-quadratic (LQ) case where the BSDE from the maximum principle leads to a Riccati equation system. It is mentioned in these papers that due to the non-markovian nature of the McKean-Vlasov systems, dynamic programming (also called Bellman optimality) principle does not hold and the problem is time inconsistent in general. Indeed, the standard Markov property of the state process, say X, is ruled out, however, as noticed in [3], this can be restored by working with the marginal distribution of X. The dynamic programming has then been applied independently in [4] and [12] for a specific control problem where the objective function depends upon statistics of X like its mean value, with a mean-field interaction on the drift of the diffusion dynamics of X, and in particular by assuming the existence at all times of a density function for the marginal distribution of X.

The purpose of this paper is to provide a detailed analysis of the dynamic programming method for the optimal control of nonlinear mean-field systems in discrete time, where the coefficients may depend both upon the marginal distributions of the state and of the control. The case of continuous time McKean-Vlasov equations requires more technicalities and mathematical tools, and will be addressed in [14]. The discrete time framework has been also considered in [9] for LQ problem, and arises naturally in situations where signal values are available only at certain times. On the other hand, it can also be viewed as the discrete time version or approximation of the optimal control of continuous time McKean-Vlasov stochastic differential equations. Our methodology is the following. By using closed-loop (also called feedback) controls, we first convert the stochastic optimal control problem into a deterministic control problem involving only the marginal distribution of the state process. We then derive the deterministic evolution of the controlled marginal distribution, and prove in its general form the dynamic programming principle (DPP). This gives sufficient conditions for optimality in terms of calculus of variations in the space of feedback control functions. Classical DPP for stochastic control problem without mean-field interaction falls within our approach. We finally apply our method for solving explicitly the mean-variance portfolio selection problem and the multivariate LQ mean-field control problem, and retrieve in particular the results obtained in [9] by four different approaches.

The outline of the paper is as follows. The next section formulates the McKean-Vlasov control problem in discrete time. In Section 3, we develop the dynamic programming method in this framework. Section 4 is devoted to applications of the DPP with explicit solutions in the LQ case including the mean-variance problem.

2 McKean-Vlasov control problem

We consider a general class of optimal control of mean-field type in discrete time. We are given two measurable spaces $(E, \mathcal{B}(E))$ and $(A, \mathcal{B}(A))$ representing respectively the state space, and the control space. We denote by $\mathcal{P}(E)$ and $\mathcal{P}(A)$ the set of probability measures on $(E, \mathcal{B}(E))$ and $(A, \mathcal{B}(A))$. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a controlled stochastic dynamics of McKean-Vlasov type:

$$X_{k+1}^\alpha \;=\; F_{k+1}(X_k^\alpha, \mathbb{P}_{X_k^\alpha}, \alpha_k, \mathbb{P}_{\alpha_k}, \varepsilon_{k+1}), \quad k \in \mathbb{N}, \quad X_0^\alpha \;=\; \xi, \tag{2.1}$$

for some measurable functions F_k defined from $E \times \mathcal{P}(E) \times A \times \mathcal{P}(A) \times \mathbb{R}^d$ into E, where $(\varepsilon_k)_k$ is a sequence of i.i.d. random variables, independent of the initial random value ξ, and we denote by $\mathbb{F} = (\mathcal{F}_k)_k$ the filtration with $\mathcal{F}_k$ the σ-algebra generated by $\{\xi, \varepsilon_1, \ldots, \varepsilon_k\}$. Here, $(X_k^\alpha)_k$ is the state process valued in E controlled by the $\mathbb{F}$-adapted process $(\alpha_k)_k$ valued in A, and we adopted the usual notation in the sequel of the paper: given a random variable Y on $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{P}_{_Y}$ denotes its probability distribution under $\mathbb{P}$. Thus, the dynamics of (X_k^α) depends at any time k of its marginal distribution, but also of the marginal distribution of the control, which represents an additional mean-field feature with respect to classical McKean Vlasov equations, and also considered recently in [9] and [16].

Let us now precise the assumptions on the McKean-Vlasov equation. We shall assume that $(E, |.|)$ is a normed space (most often $\mathbb{R}^d$), $(A, |.|)$ is also a normed space (typically a subset of $\mathbb{R}^m$), and we denote by $\mathcal{P}_2(E)$ the space of square integrable probability measures over E , i.e. $\mu \in \mathcal{P}(E)$ s.t. $\|\mu\|_2^2 := \int_E |x|^2 \mu(dx) < \infty$, and similarly for $\mathcal{P}_2(A)$. For any $(x, \mu, a, \lambda) \in E \times \mathcal{P}(E) \times A \times \mathcal{P}(A)$, and $k \in \mathbb{N}$, we denote by $P_{k+1}(x, \mu, a, \lambda, dx')$ the probability distribution of the E-valued random variable $F_{k+1}(x, \mu, a, \lambda, \varepsilon_{k+1})$ on $(\Omega, \mathcal{F}, \mathbb{P})$, and we assume

(H1) For any $k \in \mathbb{N}$, there exists some positive constant $C_{k,F}$ s.t. for all $(x, a, \mu, \lambda) \in E \times A \times \mathcal{P}(E) \times \mathcal{P}(A)$:

$$\begin{aligned} \int_E |x'|^2 P_{k+1}(x, \mu, a, \lambda, dx') &= \mathbb{E}\Big[\big|F_{k+1}(x, \mu, a, \lambda, \varepsilon_{k+1})\big|^2\Big] \\ &\leq C_{k,F}(1 + |x|^2 + |a|^2 + \|\mu\|_2^2 + \|\lambda\|_2^2). \end{aligned}$$

Assuming that the initial random value ξ is square integrable, and considering admissible controls α which are square integrable, i.e. $\mathbb{E}|\alpha_k|^2 < \infty$, for any k, it is then clear under **(H1)** that $\mathbb{E}|X_k^\alpha|^2 < \infty$, i.e. $\mathbb{P}_{_{X_k^\alpha}} \in \mathcal{P}_2(E)$, and there exists some positive constant C_k s.t.

$$\mathbb{E}|X_k^\alpha|^2 \;\leq\; C_k\big(1 + \mathbb{E}|\xi|^2 + \sum_{j=0}^{k-1} \mathbb{E}|\alpha_j|^2\big). \tag{2.2}$$

The cost functional associated to the system (2.1) over a finite horizon $n \in \mathbb{N} \setminus \{0\}$ is:

$$J(\alpha) \;:=\; \mathbb{E}\Big[\sum_{k=0}^{n-1} f_k(X_k^\alpha, \mathbb{P}_{_{X_k^\alpha}}, \alpha_k, \mathbb{P}_{_{\alpha_k}}) \;+\; g(X_n^\alpha, \mathbb{P}_{_{X_n^\alpha}})\Big], \tag{2.3}$$

for any square integrable $\mathbb{F}$-adapted processes α valued in A, where the running cost functions f_k, $k = 0, \ldots, n-1$, are measurable real-valued functions on $E \times \mathcal{P}_2(E) \times A \times \mathcal{P}_2(A)$, and the terminal cost function g is a real-valued measurable function on $E \times \mathcal{P}_2(E)$. We shall assume

(H2) There exist some positive constant C_g and for any $k = 0, \ldots, n-1$, some positive constant $C_{k,f}$ s.t. for all $(x, a, \mu, \lambda) \in E \times A \times \mathcal{P}_2(E) \times \mathcal{P}_2(A)$:

$$\begin{aligned} \big|f_k(x, \mu, a, \lambda)\big| &\leq C_{k,f}(1 + |x|^2 + |a|^2 + \|\mu\|_2^2 + \|\lambda\|_2^2), \\ \big|g(x, \mu)\big| &\leq C_g(1 + |x|^2 + \|\mu\|_2^2). \end{aligned}$$

Under **(H1)**-**(H2)**, the cost functional $J(\alpha)$ is well-defined and finite for any admissible control, and the objective is to minimize over all admissible controls the cost functional, i.e. by solving

$$V_0 \quad := \quad \inf_{\alpha} J(\alpha), \tag{2.4}$$

and when $V_0 > -\infty$, find an optimal control α^* i.e. achieving the minimum in (2.4) if it exists.

Problem (2.1)-(2.4) arises in the study of collective behaviors of a large number of players (particles) resulting from mean-field interactions: typically, the controlled dynamics of a system of N symmetric particles are given by

$$X^{i,\alpha^i}_{k+1} \quad = \quad F_{k+1}(X^{i,\alpha^i}_k, \frac{1}{N}\sum_{j=1}^N \delta_{X^{j,\alpha^i}_k}, \alpha^i_k, \frac{1}{N}\sum_{j=1}^N \delta_{\alpha^j_k}, \varepsilon^i_{k+1}), \quad i = 1, \ldots, N,$$

(here δ_x is the Dirac measure at x) and by assuming that a center decides of the general same policy $\alpha^i = \alpha$ for all players with same running and terminal gain functions, the propagation of chaos argument from McKean-Vlasov theory (see [15]) states that when the number of players N goes to infinity, the problem of each agent is asymptotically reduced to the problem of a single agent with controlled dynamics (2.1) and objective (2.4). We refer to [8] for a detailed discussion about optimal control of McKean-Vlasov equations and connection with equilibrium of large populations of individuals with mean-field interactions.

3 Dynamic programming

In this section, we make the standing assumptions **(H1)**-**(H2)**, and our purpose is to show that dynamic programming principle holds for problem (2.4), which we would like to combine with some Markov property of the controlled state process. However, notice that the McKean-Vlasov type dependence on the dynamics of the state process rules out the standard Markov property of the controlled process (X^α_k). Actually, this Markov property can be restored by considering its probability law $(\mathbb{P}_{X^\alpha_k})_k$. To be more precise and for the sake of definiteness, we shall restrict ourselves to controls $\alpha = (\alpha_k)_k$ given in closed-loop (or feedback) form:

$$\alpha_k \quad = \quad \tilde{\alpha}_k(X^\alpha_k), \qquad k = 0, \ldots, n-1, \tag{3.1}$$

for some deterministic measurable functions $\tilde{\alpha}_k$ of the state. Notice that the feedback control may also depend on the (deterministic) marginal distribution, and it will be indeed the case for the optimal one, but to alleviate notation, we omit this dependence which is implicit through the deterministic function $\tilde{\alpha}_k$. We denote by A^E the set of measurable functions on E valued in A, which satisfy a linear growth condition, and by $\mathcal{A}$ the set of admissible controls α in closed loop form (3.1) with $\tilde{\alpha}_k$ in A^E, $k \in \mathbb{N}$. We shall often identify $\alpha \in \mathcal{A}$ with the sequence $(\tilde{\alpha}_k)_k$ in A^E via (3.1). Notice that any $\alpha \in \mathcal{A}$ satisfies the square integrability condition, i.e. $\mathbb{E}|\alpha_k|^2 < \infty$, for all k. Indeed from the linear growth condition on $\tilde{\alpha}_k$ in A^E, we have $\mathbb{E}|\alpha_k|^2 \leq C_\alpha(1 + \mathbb{E}|X^\alpha_k|^2)$ for some constant C_α (depending on α), which gives the square integrability condition by (2.2).

Next, we show that the initial stochastic control problem can be reduced to a deterministic control problem. Indeed, the key point is to observe by definition of $\mathbb{P}_{X_k^\alpha}$ and noting that $\mathbb{P}_{\alpha_k}$ is the image by $\tilde{\alpha}_k$ of $\mathbb{P}_{X_k^\alpha}$ for a feedback control $\alpha \in \mathcal{A}$, that the gain functional in (2.3) can be rewritten as:

$$J(\alpha) = \sum_{k=0}^{n-1} \hat{f}_k(\mathbb{P}_{X_k^\alpha}, \tilde{\alpha}_k) + \hat{g}(\mathbb{P}_{X_n^\alpha}), \tag{3.2}$$

where $\hat{f}_k$, $k = 0, \ldots, n-1$, are defined on $\mathcal{P}_2(E) \times A^E$, $\hat{g}$ is defined on $\mathcal{P}_2(E)$ by:

$$\hat{f}_k(\mu, \tilde{\alpha}) := \int_E f_k(x, \mu, \tilde{\alpha}(x), \tilde{\alpha} \star \mu)\mu(dx), \qquad \hat{g}(\mu) := \int_E g(x, \mu)\mu(dx), \tag{3.3}$$

and $\tilde{\alpha} \star \mu \in \mathcal{P}_2(A)$ denotes the image by $\tilde{\alpha} \in A^E$ of the measure $\mu \in \mathcal{P}_2(E)$:

$$\big(\tilde{\alpha} \star \mu\big)(B) = \mu\big(\tilde{\alpha}^{-1}(B)\big), \quad \forall B \in \mathcal{B}(A).$$

Hence, the original problem (2.4) is transformed into a deterministic control problem involving the infinite dimensional marginal distribution process. Let us then define the dynamic version for problem (2.4):

$$V_k^\alpha := \inf_{\beta \in \mathcal{A}_k(\alpha)} \sum_{j=k}^{n-1} \hat{f}_j(\mathbb{P}_{X_j^\beta}, \tilde{\beta}_j) + \hat{g}(\mathbb{P}_{X_n^\beta}), \quad k = 0, \ldots, n, \tag{3.4}$$

for $\alpha \in \mathcal{A}$, where $\mathcal{A}_k(\alpha) = \{\beta \in \mathcal{A} : \beta_j = \alpha_j, j = 0, \ldots, k-1\}$, with the convention that $\mathcal{A}_0(\alpha) = \mathcal{A}$, so that $V_0 = \inf_{\alpha \in \mathcal{A}} J(\alpha)$ is equal to V_0^α. It is clear that $V_k^\alpha < \infty$, and we shall assume that

$$V_k^\alpha > -\infty, \qquad k = 0, \ldots, n, \;\; \alpha \in \mathcal{A}. \tag{3.5}$$

Remark 3.1 The finiteness condition (3.5) can be checked a priori directly from the assumptions on the model. For example, when f_k, g, hence $\hat{f}_k$, g, $k = 0, \ldots, n-1$, are lower-bounded functions, condition (3.5) clearly holds. Another example is the case when $f_k(x, \mu, a, \lambda)$, $k = 0, \ldots, n-1$, and g are lower bounded by a quadratic function in x, μ, and λ, so that by the linear growth condition on $\tilde{\alpha}$,

$$\hat{f}_k(\mu, \tilde{\alpha}) + \hat{g}(x, \mu) \geq -C_k\big(1 + \|\mu\|_2\big), \quad \forall \mu \in \mathcal{P}_2(E), \tilde{\alpha} \in A^E,$$

and we are able to derive moment estimates on X_k^α, uniformly in α: $\big\|\mathbb{P}_{X_k^\alpha}\big\|_2^2 = \mathbb{E}[|X_k^\alpha|^2] \leq C_k$, which arises typically when A is bounded from (2.2). Then, it is clear that (3.5) holds true. Otherwise, this finiteness condition can be checked a posteriori from a verification theorem, see Theorem 3.2. $\square$

The dynamic programming principle (DPP) for the deterministic control problem (3.4) takes the following formulation:

Lemma 3.1 *(Dynamic Programming Principle)*
Under (3.5), *we have*

$$\left\{ \begin{array}{rcl} V_n^\alpha & = & \hat g(\mathbb{P}_{X_n^\alpha}) \\ V_k^\alpha & = & \inf\limits_{\beta\in\mathcal{A}_k(\alpha)} \hat f_k(\mathbb{P}_{X_k^\beta}, \tilde\beta_k) + V_{k+1}^\beta, \quad k=0,\ldots,n-1. \end{array}\right. \tag{3.6}$$

Proof. In the context of deterministic control problem, the proof of the DPP is standard and does not require any measurable selection arguments. For sake of completeness and since it is quite elementary, we give it. Denote by $J_k(\alpha)$ the cost functional at time k, i.e.

$$J_k(\alpha) \;:=\; \sum_{j=k}^{n-1} \hat f_k(\mathbb{P}_{X_k^\alpha}, \tilde\alpha_k) + \hat g(\mathbb{P}_{X_n^\alpha}), \quad k=0,\ldots,n,$$

so that $V_k^\alpha = \inf_{\beta\in\mathcal{A}_k(\alpha)} J_k(\beta)$, and by W_k^α the r.h.s. of (3.6). Then,

$$\begin{aligned} W_k^\alpha &= \inf_{\beta\in\mathcal{A}_k(\alpha)} \big[\hat f_k(\mathbb{P}_{X_k^\beta}, \tilde\beta_k) + \inf_{\gamma\in\mathcal{A}_{k+1}(\beta)} J_{k+1}(\gamma)\big] \\ &= \inf_{\beta\in\mathcal{A}_k(\alpha)} \inf_{\gamma\in\mathcal{A}_{k+1}(\beta)} \big[\hat f_k(\mathbb{P}_{X_k^\beta}, \tilde\beta_k) + J_{k+1}(\gamma)\big] \\ &= \inf_{\beta\in\mathcal{A}_k(\alpha)} \inf_{\gamma\in\mathcal{A}_{k+1}(\beta)} \big[\hat f_k(\mathbb{P}_{X_k^\gamma}, \tilde\gamma_k) + J_{k+1}(\gamma)\big] \\ &= \inf_{\gamma\in\{\mathcal{A}_{k+1}(\beta):\beta\in\mathcal{A}_k(\alpha)\}} J_k(\gamma), \end{aligned}$$

where we used in the third equality the fact that $X_k^\beta = X_k^\gamma$, $\beta_k = \gamma_k$ for $\gamma\in\mathcal{A}_{k+1}(\beta)$. Finally, we notice that $\{\mathcal{A}_{k+1}(\beta) : \beta\in\mathcal{A}_k(\alpha)\} = \mathcal{A}_k(\alpha)$. Indeed, the inclusion $\subset$ is clear while for the converse inclusion, it suffices to observe that any γ in $\mathcal{A}_k(\alpha)$ satisfies obviously $\gamma\in\mathcal{A}_{k+1}(\gamma)$. This proves the required equality: $W_k^\alpha = V_k^\alpha$. $\Box$

Let us now show how one can simplify the DPP by exploiting the flow property of $(\mathbb{P}_{X_k^\alpha})_k$ for any admissible control α in feedback form $\in\mathcal{A}$. Actually, we can derive the evolution of the controlled deterministic process $(\mathbb{P}_{X_k^\alpha})_k$.

Lemma 3.2 *For any admissible control in closed-loop form $\alpha\in\mathcal{A}$, we have*

$$\mathbb{P}_{X_{k+1}^\alpha} \;=\; \Phi_{k+1}\big(\mathbb{P}_{X_k^\alpha}, \tilde\alpha_k\big), \quad k\in\mathbb{N}, \;\; \mathbb{P}_{X_0^\alpha} = \mathbb{P}_\xi \tag{3.7}$$

where Φ_{k+1} is the measurable function defined from $\mathcal{P}_2(E)\times A^E$ into $\mathcal{P}_2(E)$ by:

$$\Phi_{k+1}(\mu,\tilde\alpha)(dx') \;=\; \int_E \mu(dx) P_{k+1}(x,\mu,\tilde\alpha(x),\tilde\alpha\star\mu,dx'). \tag{3.8}$$

Proof. Fix $\alpha\in\mathcal{A}$. Recall from the definition of the transition probability $P_{k+1}(x,\mu,a,\lambda,dx')$ associated to (2.1) that

$$\mathbb{P}\big[X_{k+1}^\alpha \in dx' \big| \mathcal{F}_k\big] \;=\; P_{k+1}(X_k^\alpha, \mathbb{P}_{X_k^\alpha}, \alpha_k, \mathbb{P}_{\alpha_k}, dx'), \quad k\in\mathbb{N}. \tag{3.9}$$

For any bounded measurable function φ on E, we have by the law of iterated conditional expectation and (3.9):

$$\begin{aligned}
\mathbb{E}\Big[\varphi(X^\alpha_{k+1})\Big] &= \mathbb{E}\Big[\mathbb{E}\big[\varphi(X^\alpha_{k+1})\big|\mathcal{F}_k\big]\Big] \\
&= \mathbb{E}\Big[\int_E \varphi(x')P_{k+1}(X^\alpha_k, \mathbb{P}_{X^\alpha_k}, \alpha_k, \mathbb{P}_{\alpha_k}, dx')\Big] \\
&= \mathbb{E}\Big[\int_{E\times E} \varphi(x')P_{k+1}(x, \mathbb{P}_{X^\alpha_k}, \tilde\alpha_k(x), \tilde\alpha_k \star \mathbb{P}_{X^\alpha_k}, dx')\mathbb{P}_{X^\alpha_k}(dx)\Big]
\end{aligned}$$

where we used in the last equality the fact that $\alpha_k = \tilde\alpha_k(X^\alpha_k)$ is in closed loop form, the definition of $\mathbb{P}_{X^\alpha_k}$, and noting that $\mathbb{P}_{\alpha_k} = \tilde\alpha_k \star \mathbb{P}_{X^\alpha_k}$. This shows the required inductive relation for $\mathbb{P}_{X^\alpha_k}$. □

Remark 3.2 Relation (3.7) is the Fokker-Planck equation in discrete time for the marginal distribution of the controlled process (X^α_k). In absence of control and McKean-Vlasov type dependence, i.e. $P_{k+1}(x, dx')$ does not depend on (μ, a, λ), we retrieve the standard Fokker-Planck equation with a linear function $\Phi_{k+1}(\mu) = \mu P_{k+1}$. In our McKean-Vlasov control context, the function $\Phi_{k+1}(\mu, \tilde\alpha)$ is nonlinear in μ. □

By exploiting the inductive relation (3.7) on the controlled process $(\mathbb{P}_{X^\alpha_k})_k$, the calculation of the value processes V^α_k can be reduced to the recursive computation of deterministic functions (called value functions) on $\mathcal{P}(E)$.

Theorem 3.1 *(Dynamic programming and value functions)*
Under (3.5), we have for any $\alpha \in \mathcal{A}$, $V^\alpha_k = v_k(\mathbb{P}_{X^\alpha_k})$, $k = 0, \ldots, n$, where $(v_k)_k$ is the sequence of value functions defined recursively on $\mathcal{P}_2(E)$ by:

$$\begin{cases} v_n(\mu) = \hat g(\mu) \\ v_k(\mu) = \inf_{\tilde\alpha \in A^E} \Big[\hat f_k(\mu, \tilde\alpha) + v_{k+1}\big(\Phi_{k+1}(\mu, \tilde\alpha)\big)\Big] \end{cases} \tag{3.10}$$

for $k = 0, \ldots, n-1$, $\mu \in \mathcal{P}_2(E)$.

Proof. First observe that for any $\beta \in \mathcal{A}_k(\alpha)$, $X^\beta_k = X^\alpha_k$, $k = 0, \ldots, n$. Let us prove the result by backward induction. For $k = n$, the result clearly holds since $V^\alpha_n = \hat g(\mathbb{P}_{X^\alpha_n})$. Suppose now that at time $k+1$, $V^\alpha_{k+1} = v_{k+1}(\mathbb{P}_{X^\alpha_{k+1}})$ for some deterministic function v_{k+1} and any $\alpha \in \mathcal{A}$. Then, from the DPP (3.6) and Lemma 3.2, we get

$$\begin{aligned}
V^\alpha_k &= \inf_{\beta \in \mathcal{A}_k(\alpha)} \hat f_k(\mathbb{P}_{X^\alpha_k}, \tilde\beta_k) + v_{k+1}(\mathbb{P}_{X^\beta_{k+1}}) \\
&= \inf_{\beta \in \mathcal{A}_k(\alpha)} w_k(\mathbb{P}_{X^\alpha_k}, \tilde\beta_k)
\end{aligned} \tag{3.11}$$

where

$$w_k(\mu, \tilde\alpha) := \hat f_k(\mu, \tilde\alpha) + v_{k+1}\big(\Phi_{k+1}(\mu, \tilde\alpha)\big).$$

Now, for any $\beta \in \mathcal{A}_k(\alpha)$, and since $\tilde\beta_k$ is valued in A^E, we clearly have: $w_k(\mu, \beta_k) \geq \inf_{\tilde\alpha \in A^E} w_k(\mu, \tilde\alpha)$, and so $\inf_{\beta \in \mathcal{A}_k(\alpha)} w_k(\mu, \tilde\beta_k) \geq \inf_{\tilde\alpha \in A^E} w_k(\mu, \tilde\alpha)$. Conversely, for any $\tilde\alpha \in$

A^E, the control β defined by $\beta_j = \alpha_j$, $j \leq k-1$, and $\tilde{\beta}_j = \tilde{\alpha}$ for $j \geq k$, lies in $\mathcal{A}_k(\alpha)$, so: $w_k(\mu, \tilde{\alpha}) \geq \inf_{\beta \in \mathcal{A}_k(\alpha)} w_k(\mu, \tilde{\beta}_k)$, and thus $\inf_{\beta \in \mathcal{A}_k(\alpha)} w_k(\mu, \tilde{\beta}_k) = \inf_{\tilde{\alpha} \in A^E} w_k(\mu, \tilde{\alpha})$. We conclude from (3.11) that: $V_k^\alpha = v_k(\mathbb{P}_{X_k^\alpha})$ with $v_k(\mu) = \inf_{\tilde{\alpha} \in A^E} w_k(\mu, \tilde{\alpha})$, i.e. given by (3.10). □

Remark 3.3 Problem (2.4) includes the case where the cost functional in (2.3) is a nonlinear function of the expected value of the state process, i.e. the running cost functions and the terminal gain function are in the form: $f_k(X_k^\alpha, \mathbb{P}_{X_k^\alpha}, \alpha_k) = \bar{f}_k(X_k^\alpha, \mathbb{E}[X_k^\alpha], \alpha_k)$, $k = 0, \ldots, n-1$, $g(X_n^\alpha, \mathbb{P}_{X_n^\alpha}) = \bar{g}(X_n^\alpha, \mathbb{E}[X_n^\alpha])$, which arise for example in mean-variance problem (see Section 4). It is claimed in [5] and [16] that Bellman optimality principle does not hold, and therefore the problem is time-inconsistent. This is true when one takes into account only the state process X^α (that is its realization), since it is not Markovian, but as shown in this section, dynamic programming principle holds whenever we consider the marginal distribution as state variable. This gives more information and the price to paid is the infinite-dimensional feature of the marginal distribution state variable. □

We complete the above Bellman's optimality principle with a verification theorem, which gives a sufficient condition for finding an optimal control.

Theorem 3.2 *(Verification theorem)*
(i) Suppose we can find a sequence of real-valued functions w_k, $k = 0, \ldots, n$, defined on $\mathcal{P}_2(E)$ and satisfying the dynamic programming relation:

$$\begin{cases} w_n(\mu) &= \hat{g}(\mu) \\ w_k(\mu) &= \inf_{\tilde{\alpha} \in A^E} \Big[\hat{f}_k(\mu, \tilde{\alpha}) + w_{k+1}\big(\Phi_{k+1}(\mu, \tilde{\alpha})\big) \Big] \end{cases} \tag{3.12}$$

for $k = 0, \ldots, n-1$, $\mu \in \mathcal{P}_2(E)$. Then $V_k^\alpha = w_k(\mathbb{P}_{X_k^\alpha})$, for all $k = 0, \ldots, n$, $\alpha \in \mathcal{A}$, and thus $w_k = v_k$.
(ii) Moreover, suppose that at any time k and $\mu \in \mathcal{P}(E)$, the infimum in (3.12) for $w_k(\mu)$ is attained, by some $\tilde{\alpha}_k^*(., \mu)$ in A^E. Then, by defining by induction the control α^* in feedback form by $\alpha_k^* = \tilde{\alpha}_k^*(X_k^{\alpha^*}, \mathbb{P}_{X_k^{\alpha^*}})$, $k = 0, \ldots, n-1$, we have

$$V_0 = J(\alpha^*),$$

which means that $\alpha^* \in \mathcal{A}$ is an optimal control.

Proof. (i) Fix some $\alpha \in \mathcal{A}$, and arbitrary $\beta \in \mathcal{A}$ associated to a feedback sequence $(\tilde{\beta}_k)_k$ in A^E. Then, from the dynamic programming relation (3.12) for w_k, and recalling the evolution (3.7) of the controlled marginal distribution $\mathbb{P}_{X_k^\beta}$, we have

$$w_k(\mathbb{P}_{X_k^\beta}) \leq \hat{f}_k(\mathbb{P}_{X_k^\beta}, \tilde{\beta}_k) + v_{k+1}(\mathbb{P}_{X_{k+1}^\beta}), \quad k = 0, \ldots, n-1.$$

By induction and since $w_n = \hat{g}$, this gives

$$w_k(\mathbb{P}_{X_k^\beta}) \leq \sum_{j=k}^{n-1} \hat{f}_j(\mathbb{P}_{X_j^\beta}, \tilde{\beta}_j) + \hat{g}(\mathbb{P}_{X_n^\beta}).$$

By noting that $\mathbb{P}_{X_k^\alpha} = \mathbb{P}_{X_k^\beta}$, when $\beta \in \mathcal{A}_k(\alpha)$, and since β is arbitrary, this proves that $w_k(\mathbb{P}_{X_k^\alpha}) \leq V_k^\alpha$. In particular, $V_k^\alpha > -\infty$, i.e. relation (3.5) holds, and then by Theorem 3.1, V_k^α is characterized by the sequence of value functions $(v_k)_k$ defined by the DP (3.10). This DP obviously defines by backward induction a unique sequence of functions on $\mathcal{P}_2(E)$, hence $w_k = v_k$, $k = 0, \ldots, n$, and therefore $V_k^\alpha = w_k(\mathbb{P}_{X_k^\alpha})$.
(ii) By definition of $\tilde{\alpha}_k^*$ which attains the infimum in (3.12), we have

$$w_k\big(\mathbb{P}_{X_k^{\alpha^*}}\big) \;=\; \hat{f}_k(\mathbb{P}_{X_k^{\alpha^*}}, \tilde{\alpha}_k^*(., \mathbb{P}_{X_k^{\alpha^*}})) + w_{k+1}\big(\mathbb{P}_{X_{k+1}^{\alpha^*}}\big), \quad k = 0, \ldots, n-1.$$

By induction this implies that

$$V_0 \;=\; w_0(\mathbb{P}_\xi) \;=\; \sum_{k=0}^{n-1} \hat{f}_k\big(\mathbb{P}_{X_k^{\alpha^*}}, \tilde{\alpha}_k^*(., \mathbb{P}_{X_k^{\alpha^*}})\big) \;+\; \hat{g}(\mathbb{P}_{X_n^{\alpha^*}}) \;=\; J(\alpha^*),$$

which shows that α^* is an optimal control. □

The above verification theorem, which consists in solving the dynamic programming relation (3.12), is useful to check a posteriori the finiteness condition (3.5), and can be applied in practice fo find explicit solutions to some McKean-Vlasov control problems, as investigated in the next section.

4 Applications

4.1 Special cases

We consider some particular cases, and provide the special forms of the DPP.

4.1.1 No mean-field interaction

We first consider the standard control case where there is no mean-field interaction in the dynamics of the state process, i.e. $F_{k+1}(x, a, \varepsilon_{k+1})$, hence $P_{k+1}(x, a, dx')$ do not depend on μ, λ, as well as in the cost functions $f_k(x, a)$ and $g(x)$. For simplicity, we assume that A is a bounded set, which ensures the finiteness condition (3.5). In this case, we can see that the value functions v_k are in the form

$$v_k(\mu) \;=\; \int_E \tilde{v}_k(x)\mu(dx), \quad k = 0, \ldots, n, \tag{4.1}$$

where the functions $\tilde{v}_k$ defined on E satisfy the classical dynamic programming principle:

$$\begin{cases} \tilde{v}_n(x) \;=\; g(x) \\ \tilde{v}_k(x) \;=\; \inf_{a \in A} \Big[f_k(x, a) + \mathbb{E}\big[\tilde{v}_{k+1}(X_{k+1}^\alpha)\big| X_k^\alpha = x, \alpha_k = a\big]\Big], \end{cases} \tag{4.2}$$

for $k = 0, \ldots, n-1$. Let us check this result by backward induction. This holds true for $k = n$ since $v_n(\mu) = \hat{g}(\mu) = \int_E g(x)\mu(dx)$. Suppose that (4.1) holds true at time $k+1$. Then, from the DPP (3.10), (3.8) and Fubini's theorem, we have

$$
\begin{aligned}
v_k(\mu) &= \inf_{\tilde\alpha\in A^E}\Big[\int_E f_k(x,\tilde\alpha(x))\mu(dx) + \int_E \tilde v_{k+1}(x')\Phi_{k+1}(\mu,\tilde\alpha)(dx')\Big] \\
&= \inf_{\tilde\alpha\in A^E}\Big[\int_E \big[f_k(x,\tilde\alpha(x)) + \int_E \tilde v_{k+1}(x')P_{k+1}(x,\tilde\alpha(x),dx')\big]\mu(dx)\Big] \\
&= \inf_{\tilde\alpha\in A^E}\int_E \tilde w_k(x,\tilde\alpha(x))\mu(dx)
\end{aligned}
$$

where we set $\tilde w_k(x,a) = f_k(x,a) + \int_E \tilde v_{k+1}(x')P_{k+1}(x,a,dx')$. Now, we observe that

$$
\inf_{\tilde\alpha\in A^E}\int_E \tilde w_k(x,\tilde\alpha(x))\mu(dx) = \int_E \inf_{a\in A}\tilde w_k(x,a)\mu(dx). \tag{4.3}
$$

Indeed, since for any $\tilde\alpha \in A^E$, the value $\tilde\alpha(x)$ is valued in A for any $x \in E$, it is clear that the inequality $\geq$ in (4.3) holds true. Conversely, for any $\varepsilon > 0$, and $x \in E$, one can find $\tilde\alpha^\varepsilon(x)$ in A such that

$$
\tilde w_k(x,\tilde\alpha^\varepsilon(x)) \leq \inf_{a\in A}\tilde w_k(x,a) + \varepsilon.
$$

By a measurable selection theorem, the map $x \mapsto \tilde\alpha^\varepsilon(x)$ can be chosen measurable, and since A is bounded, the function $\tilde\alpha^\varepsilon$ lies in A^E. It follows that

$$
\inf_{\tilde\alpha\in A^E}\int_E \tilde w_k(x,\tilde\alpha(x))\mu(dx) \leq \int_E \tilde w_k(x,\tilde\alpha^\varepsilon(x))\mu(dx) \leq \int_E \inf_{a\in A}\tilde w_k(x,a)\mu(dx) + \varepsilon,
$$

which shows (4.3) since ε is arbitrary. Therefore, we have $v_k(\mu) = \int_E \tilde v_k(x)\mu(dx)$ with

$$
\begin{aligned}
\tilde v_k(x) &= \inf_{a\in A}\tilde w_k(x,a) \\
&= \inf_{a\in A}\big[f_k(x,a) + \int_E \tilde v_{k+1}(x')P_{k+1}(x,a,dx')\big],
\end{aligned}
$$

which is the relation (4.2) at time k from the definition of the transition probability P_{k+1}.

4.1.2 First order interactions

We consider the case of first order interactions, i.e. the dependence of the model coefficients upon the probability measures is linear in the sense that for any $(x,\mu,a) \in E\times\mathcal{P}_2(E)\times A$, $\tilde\alpha \in A^E$,

$$
\begin{aligned}
P_{k+1}(x,\mu,a,\tilde\alpha\star\mu,dx') &= \int_E \tilde P_{k+1}(x,y,a,\tilde\alpha(y),dx')\mu(dy), \\
f_k(x,\mu,a,\tilde\alpha\star\mu) &= \int_E \tilde f_k(x,y,a,\tilde\alpha(y))\mu(dy), \qquad g(x,\mu) = \int_E \tilde g(x,y)\mu(dy),
\end{aligned}
$$

for some transition probability kernels $\tilde P_{k+1}$ from $E\times E\times A\times A$ into E, measurable functions $\tilde f_k$ defined on $E\times E\times A\times A$, $k = 0,\ldots,n-1$, and $\tilde g$ defined on $E\times E$. In this case, the value functions v_k are in the reduced form

$$
v_k(\mu) = \int_{E^{2^{n-k+1}}} \tilde v_k(\mathbf{x}_{2^{n-k+1}})\mu(d\mathbf{x}_{2^{n-k+1}}), \quad k = 0,\ldots,n,
$$

where we denote by $\mathbf{x}_p$ the p-tuple $(x_1,\ldots,x_p) \in E^p$, by $\mu(d\mathbf{x}_p)$ the product measure $\mu(dx_1)\otimes\ldots\otimes\mu(dx_p)$, and the functions $\tilde{v}_k$ are defined recursively on $E^{2^{n-k+1}}$ by

$$\begin{cases} \tilde{v}_n(x,y) & = \tilde{g}(x,y) \\ \tilde{v}_k(\mathbf{x}_{2^{n-k}},\mathbf{y}_{2^{n-k}}) & = \inf_{\tilde{\alpha}\in A^E}\Big[\tilde{f}_k(x_1,y_1,\tilde{\alpha}(x_1),\tilde{\alpha}(y_1)) \\ & \quad + \int_{E^{2^{n-k}}} \tilde{v}_{k+1}(\mathbf{x}'_{2^{n-k}})\tilde{\mathbf{P}}_{\mathbf{k+1}}(\mathbf{x_{2^{n-k}}},\mathbf{y_{2^{n-k}}},\tilde{\alpha}(\mathbf{x_{2^{n-k}}}),\tilde{\alpha}(\mathbf{y_{2^{n-k}}}),\mathbf{dx'_{2^{n-k}}})\Big], \end{cases}$$

where we set

$$\begin{aligned} & \tilde{\mathbf{P}}_{\mathbf{k+1}}(\mathbf{x_p},\mathbf{y_p},\tilde{\alpha}(\mathbf{x_p}),\tilde{\alpha}(\mathbf{y_p}),\mathbf{dx'_p}) \\ = \ & \tilde{P}_{k+1}(x_1,y_1,\tilde{\alpha}(x_1),\tilde{\alpha}(y_1),dx'_1)\otimes\ldots\otimes\tilde{P}_{k+1}(x_p,y_p,\tilde{\alpha}(x_p),\tilde{\alpha}(y_p),dx'_p). \end{aligned}$$

This result is easily checked by induction from the DPP (3.10), and it is left to the reader.

4.2 Linear-quadratic McKean-Vlasov control problem

We consider a general multivariate linear McKean-Vlasov dynamics in $E=\mathbb{R}^d$ with control valued in $A=\mathbb{R}^m$:

$$\begin{aligned} X^\alpha_{k+1} \ = \ & \big(B_kX^\alpha_k+\bar{B}_k\mathbb{E}[X^\alpha_k]+C_k\alpha_k+\bar{C}_k\mathbb{E}[\alpha_k]\big) \\ & + \big(D_kX^\alpha_k+\bar{D}_k\mathbb{E}[X^\alpha_k]+H_k\alpha_k+\bar{H}_k\mathbb{E}[\alpha_k]\big)\varepsilon_{k+1}, \quad k=0,\ldots,n-1, \end{aligned} \tag{4.4}$$

starting from $X^\alpha_0=\xi$, where B_k, $\bar{B}_k$, D_k, $\bar{D}_k$ are constant matrices in $\mathbb{R}^{d\times d}$, C_k, $\bar{C}_k$, H_k, $\bar{H}_k$ are constant matrices in $\mathbb{R}^{d\times m}$, and (ε_k) is a sequence of i.i.d. random variables with distribution $\mathcal{N}(0,1)$, independent of ξ. The quadratic cost functional to be minimized is given by

$$\begin{aligned} J(\alpha) \ = \ \mathbb{E}\Big[\sum_{k=0}^{n-1} & \big[(X^\alpha_k)^\intercal Q_kX^\alpha_k+\big(\mathbb{E}[X^\alpha_k]\big)^\intercal\bar{Q}_k\mathbb{E}[X^\alpha_k]+L_k^\intercal X^\alpha_k+\bar{L}_k^\intercal\mathbb{E}[X^\alpha_k] \\ & + \alpha_k^\intercal R_k\alpha_k+\big(\mathbb{E}[\alpha_k]\big)^\intercal\bar{R}_k\mathbb{E}[\alpha_k]\big] \\ & + \big(X^\alpha_n\big)^\intercal QX^\alpha_n+\big(\mathbb{E}[X^\alpha_n]\big)^\intercal\bar{Q}\mathbb{E}[X^\alpha_n]+L^\intercal X^\alpha_n+\bar{L}^\intercal\mathbb{E}[X^\alpha_k]\Big], \end{aligned} \tag{4.5}$$

for some constants matrices Q_k, $\bar{Q}_k$, Q, $\bar{Q}$, in $\mathbb{R}^{d\times d}$, R_k, $\bar{R}_k$ in $\mathbb{R}^{m\times m}$, and vectors L_k, $\bar{L}_k$, L, $\bar{L}\in\mathbb{R}^d$, $k=0,\ldots,n-1$. Here $x^\intercal$ denotes the transpose of a matrix/vector x. Since the cost functions are real-valued, we may assume w.l.o.g. that all these matrices Q_k, $\bar{Q}_k$, Q, $\bar{Q}$, R_k and $\bar{R}_k$ are symmetric. This model is in the form (2.1) and associated to a transition probability satisfying:

$$\begin{aligned} P_{k+1}(x,\mu,a,\lambda,dx') & \rightsquigarrow \mathcal{N}\Big(M_k(x,\mu,a,\lambda);\Sigma_k(x,\mu,a,\lambda)\Sigma_k(x,\mu,a,\lambda)^\intercal\Big) \\ M_k(x,\mu,a,\lambda) & = B_kx+\bar{B}_k\bar{\mu}+C_ka+\bar{C}_k\bar{\lambda} \\ \Sigma_k(x,\mu,a,\lambda) & = D_kx+\bar{D}_k\bar{\mu}+H_ka+\bar{H}_k\bar{\lambda} \end{aligned} \tag{4.6}$$

where we set for any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$ (resp. $\mathcal{P}_2(\mathbb{R}^m)$) symmetric matrix $\Lambda\in\mathbb{R}^{d\times d}$ (resp. in $\mathbb{R}^{m\times m}$):

$$\bar{\mu} \ := \int x\mu(dx), \quad \bar{\mu}_2(\Lambda) \ := \int x^\intercal\Lambda x\mu(dx), \quad \mathrm{Var}(\mu)(\Lambda) \ := \ \bar{\mu}_2(\Lambda)-\bar{\mu}^\intercal\Lambda\bar{\mu},$$

and in the form (2.3), hence (3.2) for feedback controls, with

$$\begin{aligned}
\hat{f}_k(\mu,\tilde{\alpha}) &= \mathrm{Var}(\mu)(Q_k) + \bar{\mu}^\intercal (Q_k+\bar{Q}_k)\bar{\mu} + (L_k+\bar{L}_k)^\intercal \bar{\mu} \\
&\quad + \mathrm{Var}(\tilde{\alpha}\star\mu)(R_k) + \overline{\tilde{\alpha}\star\mu}^\intercal (R_k+\bar{R}_k)\overline{\tilde{\alpha}\star\mu} \\
\hat{g}(\mu) &= \mathrm{Var}(\mu)(Q) \ + \ \bar{\mu}^\intercal (Q+\bar{Q})\bar{\mu} \ + \ (L+\bar{L})^\intercal \bar{\mu}.
\end{aligned}$$

We look for candidate w_k, $k=0,\ldots,n$, of values functions satisfying the dynamic programming principle (3.10), in the quadratic form:

$$w_k(\mu) \;=\; \mathrm{Var}(\mu)(\Lambda_k) + \bar{\mu}^\intercal \Gamma_k \bar{\mu} + \rho_k^\intercal \bar{\mu} + \chi_k, \tag{4.7}$$

for some constant symmetric matrices Λ_k and Γ_k in $\mathbb{R}^{d\times d}$, vector $\rho_k\in\mathbb{R}^d$ and real χ_k to be determined below. We proceed by backward induction. For $k=n$, we see that $w_k=\hat{g}$ $(=v_k)$ iff

$$\Lambda_n \;=\; Q, \qquad \Gamma_n \;=\; Q+\bar{Q}, \ \ \rho_n \;=\; L+\bar{L}, \ \chi_n \;=\; 0. \tag{4.8}$$

Now, suppose that the form (4.7) holds true at time $k+1$, and observe from (3.8) and (4.6) that for any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, $\tilde{\alpha}\in A^E$, $\Lambda\in\mathbb{R}^{d\times d}$, we have by Fubini's theorem:

$$\begin{aligned}
\overline{\Phi_{k+1}(\mu,\tilde{\alpha})} &= \int_{\mathbb{R}^d} \mathbb{E}\big[Y(x,\mu,\tilde{\alpha})\big]\mu(dx) \\
\overline{\Phi_{k+1}(\mu,\tilde{\alpha})}_2(\Lambda) &= \int_{\mathbb{R}^d} \mathbb{E}\big[Y(x,\mu,\tilde{\alpha})^\intercal \Lambda Y(x,\mu,\tilde{\alpha})\big]\mu(dx),
\end{aligned}$$

where $Y(x,\mu,\tilde{\alpha}) \rightsquigarrow \mathcal{N}\Big(M_k(x,\mu,\tilde{\alpha}(x),\tilde{\alpha}\star\mu);\Sigma_k(x,\mu,\tilde{\alpha}(x),\tilde{\alpha}\star\mu)\Sigma_k(x,\mu,\tilde{\alpha}(x),\tilde{\alpha}\star\mu)^\intercal\Big)$. Therefore,

$$\overline{\Phi_{k+1}(\mu,\tilde{\alpha})} \;=\; (B_k+\bar{B}_k)\bar{\mu} + (C_k+\bar{C}_k)\overline{\tilde{\alpha}\star\mu},$$

and after some tedious but straightforward calculation:

$$\begin{aligned}
\mathrm{Var}(\Phi_{k+1}(\mu,\tilde{\alpha}))(\Lambda) &= \overline{\Phi_{k+1}(\mu,\tilde{\alpha})}_2(\Lambda) - \overline{\Phi_{k+1}(\mu,\tilde{\alpha})}^\intercal \Lambda \overline{\Phi_{k+1}(\mu,\tilde{\alpha})} \\
&= \int_{\mathbb{R}^d} \big[\Sigma_k(x,\mu,\tilde{\alpha}(x),\tilde{\alpha}\star\mu)^\intercal \Lambda \Sigma(x,\mu,\tilde{\alpha}(x),\tilde{\alpha}\star\mu) \\
&\qquad + \ M_k(x,\mu,\tilde{\alpha}(x),\tilde{\alpha}\star\mu)^\intercal \Lambda M(x,\mu,\tilde{\alpha}(x),\tilde{\alpha}\star\mu)\big]\mu(dx) \\
&\quad - \Big((B_k+\bar{B}_k)\bar{\mu} + (C_k+\bar{C}_k)\overline{\tilde{\alpha}\star\mu}\Big)^\intercal \Lambda \Big((B_k+\bar{B}_k)\bar{\mu} + (C_k+\bar{C}_k)\overline{\tilde{\alpha}\star\mu}\Big) \\
&= \mathrm{Var}(\mu)(B_k^\intercal \Lambda B_k + D_k^\intercal \Lambda D_k) + \bar{\mu}^\intercal (D_k+\bar{D}_k)^\intercal \Lambda (D_k+\bar{D}_k)\bar{\mu} \\
&\quad + \mathrm{Var}(\tilde{\alpha}*\mu)(H_k^\intercal \Lambda H_k + C_k^\intercal \Lambda C_k) \\
&\quad + \overline{\tilde{\alpha}\star\mu}^\intercal (H_k+\bar{H}_k)^\intercal \Lambda (H_k+\bar{H}_k)\overline{\tilde{\alpha}\star\mu} \\
&\quad + \ 2\int_{\mathbb{R}^d} (x-\bar{\mu})^\intercal (D_k^\intercal \Lambda H_k + B_k^\intercal \Lambda C_k)\tilde{\alpha}(x)\mu(dx) \\
&\quad + \ 2\bar{\mu}^\intercal (D_k+\bar{D}_k)^\intercal \Lambda (H_k+\bar{H}_k)\overline{\tilde{\alpha}\star\mu}.
\end{aligned}$$

Then, w_k satisfies the DPP (3.10) iff

$$\begin{aligned} w_k(\mu) &= \inf_{\tilde\alpha \in A^E} \Big[\hat f_k(\mu,\tilde\alpha) + \mathrm{Var}(\Phi_{k+1}(\mu,\tilde\alpha))(\Lambda_{k+1}) + \overline{\Phi_{k+1}(\mu,\tilde\alpha)}^{\intercal} \Gamma_{k+1} \overline{\Phi_{k+1}(\mu,\tilde\alpha)} \Big] \quad (4.9) \\ &= \mathrm{Var}(\mu)(Q_k + B_k^{\intercal}\Lambda_{k+1}B_k + D_k^{\intercal}\Lambda_{k+1}D_k) + \inf_{\tilde\alpha\in A^E} G^{\mu}_{k+1}(\tilde\alpha) \\ &\quad + \bar\mu^{\intercal}\big(Q_k + \bar Q_k + (D_k+\bar D_k)^{\intercal}\Lambda_{k+1}(D_k+\bar D_k) + (B_k+\bar B_k)^{\intercal}\Gamma_{k+1}(B_k+\bar B_k)\big)\bar\mu \\ &\quad + \big(L_k + \bar L_k + (B_k+\bar B_k)^{\intercal}\rho_{k+1}\big)^{\intercal}\bar\mu + \chi_{k+1}, \qquad (4.10) \end{aligned}$$

where we define the function $G^{\mu}_{k+1} : L^2(\mu;A) \mapsto \mathbb{R}$ by

$$\begin{aligned} G^{\mu}_{k+1}(\tilde\alpha) &= \mathrm{Var}(\tilde\alpha \star \mu)(V_k) \ + \ \overline{\tilde\alpha\star\mu}^{\intercal} W_k \overline{\tilde\alpha\star\mu} \ + \ 2\int_{\mathbb{R}^d} (x-\bar\mu)^{\intercal} S_k \tilde\alpha(x)\mu(dx) \\ &\quad + 2\bar\mu^{\intercal} T_k \overline{\tilde\alpha\star\mu} \ + \ \rho_{k+1}^{\intercal}(C_k + \bar C_k)\overline{\tilde\alpha\star\mu}, \qquad (4.11) \end{aligned}$$

and we set $V_k = V_k(\Lambda_{k+1})$, $W_k = W_k(\Lambda_{k+1},\Gamma_{k+1})$, $S_k = S_k(\Lambda_{k+1})$, $T_k = T_k(\Lambda_{k+1},\Gamma_{k+1})$, with

$$\begin{cases} V_k(\Lambda_{k+1}) = R_k + H_k^{\intercal}\Lambda_{k+1}H_k + C_k^{\intercal}\Lambda_{k+1}C_k; \\ W_k(\Lambda_{k+1},\Gamma_{k+1}) = R_k + \bar R_k + (C_k+\bar C_k)^{\intercal}\Gamma_{k+1}(C_k+\bar C_k) + (H_k+\bar H_k)^{\intercal}\Lambda_{k+1}(H_k+\bar H_k) \\ S_k(\Lambda_{k+1}) = D_k^{\intercal}\Lambda_{k+1}H_k + B_k^{\intercal}\Lambda_{k+1}C_k; \\ T_k(\Lambda_{k+1},\Gamma_{k+1}) = (D_k+\bar D_k)^{\intercal}\Lambda_{k+1}(H_k+\bar H_k) + (B_k+\bar B_k))^{\intercal}\Gamma_{k+1}(C_k+\bar C_k). \end{cases} \tag{4.12}$$

Here, $L^2(\mu;A) \supset A^E$ is the Hilbert space of measurable functions on $E = \mathbb{R}^d$ valued in $A = \mathbb{R}^m$ and square integrable w.r.t. $\mu \in \mathcal{P}_2(E)$.

We now search for the infimum of the function G^{μ}_{k+1}, and shall make the following assumptions on the symmetric matrices of the quadratic cost functional and on the coefficients of the state dynamics:

(c0)

$$\begin{cases} Q \ \geq\ 0,\ Q + \bar Q \ \geq\ 0, \qquad Q_k \ \geq\ 0,\ Q_k + \bar Q_k \ \geq\ 0, \\ \qquad\qquad\qquad\qquad\qquad R_k \ \geq\ 0,\ R_k + \bar R_k \ \geq\ 0, \quad k = 0,\ldots,n-1, \end{cases}$$

and for all $k = 0,\ldots,n-1$ (with the convention that $Q_n = Q$, $\bar Q_n = \bar Q$)

(c1) $R_k > 0$, or $[\mathrm{rank}(C_k) = d,\ Q_{k+1} > 0]$, or $[\mathrm{rank}(H_k) = d,\ Q_{k+1} > 0]$,

(c2) $R_k + \bar R_k > 0$, or $[\mathrm{rank}(C_k + \bar C_k) = d,\ Q_{k+1} + \bar Q_{k+1} > 0]$, or $[\mathrm{rank}(H_k + \bar H_k) = d,$ $Q_{k+1} > 0]$.

Conditions **(c0)**-**(c1)**-**(c2)** is slightly weaker than the condition in [9] (see their Theorem 3.1), where the condition **(c0)** is strengthened to $R_k > 0$ and $R_k + \bar R_k > 0$ for all $k = 0,\ldots,n-1$, for ensuring the existence of an optimal control. We relax this positivity condition with the conditions **(c1)**-**(c2)** in order to include the case of mean-variance problem (see the example at the end of this section). Actually, as we shall see in Remark 4.1, these conditions will guarantee that for Λ_k, Γ_k to be determined below, the function G^{μ}_{k+1} is convex and coercive on $L^2(\mu;A)$ for any $k = 0,\ldots,n-1$. For the moment, we derive after some straightforward calculation the Gateaux derivative of G^{μ}_{k+1} at $\tilde{\alpha}$ in the direction $\beta \in L^2(\mu; A)$, which is given by:

$$DG^{\mu}_{k+1}(\tilde{\alpha};\beta) \;:=\; \lim_{\varepsilon\to 0} \frac{G^{\mu}_{k+1}(\tilde{\alpha}+\varepsilon\beta) - G^{\mu}_{k+1}(\tilde{\alpha})}{\varepsilon} \;=\; \int_{\mathbb{R}^d} g_{k+1}(x,\tilde{\alpha})\beta(x)\mu(dx)$$

with

$$\begin{aligned} g_{k+1}(x,\tilde{\alpha}) \;&=\; 2\tilde{\alpha}(x)^{\intercal}V_k + 2\overline{\tilde{\alpha}\star\mu}^{\intercal}\big(W_k - V_k\big) \\ &\quad + 2(x-\mu)^{\intercal}S_k + 2\bar{\mu}^{\intercal}T_k + \rho^{\intercal}_{k+1}(C_k + \bar{C}_k). \end{aligned}$$

We shall check later in Remark 4.1 that V_k and W_k are positive symmetric matrices, hence invertible. We thus see that $DG^{\mu}_{k+1}(\tilde{\alpha};.)$ vanishes for $\tilde{\alpha} = \tilde{\alpha}^*_k(.,\mu)$ s.t. $g_{k+1}(x,\tilde{\alpha}^*_k(.,\mu)) = 0$ for all $x \in \mathbb{R}^d$, which gives:

$$\tilde{\alpha}^*_k(x,\mu) \;=\; -V_k^{-1}S_k^{\intercal}(x-\bar{\mu}) \;-\; W_k^{-1}T_k^{\intercal}\bar{\mu} \;-\; \frac{1}{2}W_k^{-1}(C_k+\bar{C}_k)^{\intercal}\rho_{k+1} \tag{4.13}$$

and then after some straightforward caculation:

$$\begin{aligned} G^{\mu}_{k+1}(\tilde{\alpha}^*_k(.,\mu)) \;&=\; -\mathrm{Var}(\mu)\big(S_kV_k^{-1}S_k^{\intercal}\big) \;-\; \bar{\mu}^{\intercal}\big(T_kW_k^{-1}T_k^{\intercal}\big)\bar{\mu} \;-\; \bar{\mu}^{\intercal}T_kW_k^{-1}(C_k+\bar{C}_k)^{\intercal}\rho_{k+1} \\ &\quad - \frac{1}{4}\rho^{\intercal}_{k+1}(C_k+\bar{C}_k)W_k^{-1}(C_k+\bar{C}_k)^{\intercal}\rho_{k+1}. \end{aligned}$$

Assuming for the moment that $\tilde{\alpha}^*_k(.,\mu)$ attains the infimum of G^{μ}_{k+1} (this is a consequence of the convexity and coercivity of G^{μ}_{k+1} shown in Remark 4.1), and plugging the above expression in (4.9), we see that w_k is like the function $\mu \mapsto G^{\mu}_{k+1}(\tilde{\alpha}^*_k(.,\mu))$, a linear combination of terms in $\mathrm{Var}(\mu)(.)$, $\bar{\mu}^{\intercal}(.)\bar{\mu}$, and by identification with the form (4.7), we obtain an inductive relation for Λ_k, Γ_k, ρ_k, χ_k:

$$\left\{\begin{array}{lcl} \Lambda_k &=& Q_k + B_k^{\intercal}\Lambda_{k+1}B_k + D_k^{\intercal}\Lambda_{k+1}D_k - S_k(\Lambda_{k+1})V_k^{-1}(\Lambda_{k+1})S_k^{\intercal}(\Lambda_{k+1}) \\ \Gamma_k &=& (Q_k+\bar{Q}_k) + (B_k+\bar{B}_k)^{\intercal}\Gamma_{k+1}(B_k+\bar{B}_k) + (D_k+\bar{D}_k)^{\intercal}\Lambda_{k+1}(D_k+\bar{D}_k) \\ && -\, T_k(\Lambda_{k+1},\Gamma_{k+1})W_k^{-1}(\Lambda_{k+1},\Gamma_{k+1})T_k^{\intercal}(\Lambda_{k+1},\Gamma_{k+1}) \\ \rho_k &=& L_k + \bar{L}_k + \big[(B_k+\bar{B}_k) - (C_k+\bar{C}_k)W_k^{-1}(\Lambda_{k+1},\Gamma_{k+1})T_k^{\intercal}(\Lambda_{k+1},\Gamma_{k+1})\big]\rho_{k+1} \\ \chi_k &=& \chi_{k+1} - \frac{1}{4}\rho^{\intercal}_{k+1}(C_k+\bar{C}_k)W_k^{-1}(\Lambda_{k+1},\Gamma_{k+1})(C_k+\bar{C}_k)^{\intercal}\rho_{k+1}. \end{array}\right. \tag{4.14}$$

for all $k = 0,\ldots,n-1$, starting from the terminal condition (4.8). The relations for (Λ_k,Γ_k) in (4.14) are two algebraic Riccati difference equations, while the equations for ρ_k and χ_k are linear equations once (Λ_k,Γ_k) are determined. This system (4.14) is the same as the one obtained in [9]. In the particular mean-variance problem considered at the end of this section, we can obtain explicit closed-form expressions for the solutions $(\Lambda_k,\Gamma_k,\rho_k,\chi_k)$ to this Riccati system. However, in general, there are no closed-form formulae, and these quantities are simply computed by induction.

In the following remark, we check the issues that have left open up to now.

Remark 4.1 Let conditions **(c0)-(c1)-(c2)** hold. We prove by backward induction that for all $k = 1,\ldots,n$, the matrices $V_{k-1} = V_{k-1}(\Lambda_k)$, $W_{k-1} = W_{k-1}(\Lambda_k,\Gamma_k)$ are symmetric positive, hence invertible, with (Λ_k,Γ_k) given by (4.14), together with the nonnegativity of the symmetric matrices Λ_k, Γ_k, which will immediately gives the convexity and coercivity of the function G_k^μ in (4.11) for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

At time $k = n$, we have $\Lambda_n = Q \geq 0$, $\Gamma_n = Q + \bar{Q} \geq 0$, and thus from (4.12), the $d \times d$ matrices $V_{n-1} = V_{n-1}(\Lambda_n)$, $W_{n-1} = W_{n-1}(\Lambda_n, \Gamma_n)$ are symmetric positive under **(c0)-(c1)-(c2)**: indeed, for example if $R_{n-1} = 0$, then the rank condition on C_{n-1} or H_{n-1} together with the positivity of $\Lambda_n = Q$ in **(c1)** will ensure that V_{n-1} is positive. Now, suppose that the assertion is true at time $k+1$, i.e. V_k, W_k are symmetric positive, and Λ_{k+1}, Γ_{k+1} are symmetric nonnegative. Then, it is clear from (4.14) that Λ_k and Γ_k are symmetric, and noting that they can be rewritten from the expression of V_k, W_k, S_k, T_k in (4.12) as

$$\left\{ \begin{array}{rcl} \Lambda_k & = & Q_k + S_k V_k^{-1} R_k \big(S_k V_k^{-1}\big)^{\intercal} \; + \; \Big(B_k - C_k (S_k V_k^{-1})^{\intercal}\big]^{\intercal} \Lambda_{k+1} \big[B_k - C_k (S_k V_k^{-1})^{\intercal}\Big) \\ & & + \Big(D_k - H_k (S_k V_k^{-1})^{\intercal}\Big)^{\intercal} \Lambda_{k+1} \Big(D_k - H_k (S_k V_k^{-1})^{\intercal}\Big) \\ \Gamma_k & = & Q_k + \bar{Q}_k \; + \; T_k W_k^{-1} (R_k + \bar{R}_k) \big(T_k W_k^{-1}\big)^{\intercal} \\ & & + \Big(B_k + \bar{B}_k - (C_k + \bar{C}_k)(T_k W_k^{-1})^{\intercal}\Big)^{\intercal} \Gamma_{k+1} \Big(B_k + \bar{B}_k - (C_k + \bar{C}_k)(T_k W_k^{-1})^{\intercal}\Big) \\ & & + \Big(D_k + \bar{D}_k - (H_k + \bar{H}_k)(T_k W_k^{-1})^{\intercal}\Big)^{\intercal} \Gamma_{k+1} \Big(D_k + \bar{D}_k - (H_k + \bar{H}_k)(T_k W_k^{-1})^{\intercal}\Big), \end{array} \right.$$

it is also clear that they are nonnegative under **(c0)**. Finally from the expression (4.12) at time $k-1$, we see that $V_{k-1} = V_{k-1}(\Lambda_k)$ and $W_{k-1} = W_{k-1}(\Lambda_k, \Gamma_k)$ are symmetric positive under **(c0)-(c1)-(c2)**, which shows the required assertion. □

In view of the above derivation and Remark 4.1, it follows that the functions w_k, $k = 0, \ldots, n$, given in the quadratic form (4.7) with $(\Lambda_k, \Gamma_k, \rho_k, \chi_k)$ as in (4.14), satisfy by construction the DPP (3.10), and by the verification theorem, this implies that the value functions are given by $v_k = w_k$, while the optimal control is given in feedback form from (4.13) by:

$$\alpha_k^* \;=\; \tilde{\alpha}_k(X_k^*, \mathbb{P}_{X_k^*}) \;=\; -V_k^{-1} S_k^{\intercal}\big(X_k^* - \mathbb{E}[X_k^*]\big) \;-\; W_k^{-1} T_k^{\intercal} \mathbb{E}[X_k^*], \tag{4.15}$$

where $X_k^* = X_k^{\alpha^*}$ is the optimal wealth process with the feedback control α^*. We retrieve the expression obtained in [9] (see e.g. their Theorem 3.1). We can push further our calculations to get an explicit form of the optimal control expressed only in terms of the state process (and not on its mean). Indeed, from the linear dynamics (4.4), we have

$$\begin{array}{rcl} \mathbb{E}[X_{k+1}^*] & = & (B_k + \bar{B}_k)\mathbb{E}[X_k^*] + (C_k + \bar{C}_k)\mathbb{E}[\alpha_k^*] \\ & = & (B_k + \bar{B}_k)\mathbb{E}[X_k^*] - (C_k + \bar{C}_k)\big(W_k^{\intercal}\big)^{-1} T_k^{\intercal} \mathbb{E}[X_k^*] \;=\; N_k \mathbb{E}[X_k^*], \end{array}$$

with $N_k = B_k + \bar{B}_k - (C_k + \bar{C}_k) W_k^{-1} T_k$, for $k = 0, \ldots, n-1$, and so by induction:

$$\mathbb{E}[X_k^*] \;=\; N_{k-1} \ldots N_0 \mathbb{E}[\xi].$$

Plugging into (4.15), this gives the explicit form of the optimal control as

$$\alpha_k^* \;=\; -V_k^{-1} S_k^{\intercal} X_k^* \;+\; \big(V_k^{-1} S_k^{\intercal} - W_k^{-1} T_k^{\intercal}\big) N_{k-1} \ldots N_0 \mathbb{E}[\xi], \quad k = 0, \ldots, n-1. \tag{4.16}$$

We observe that the optimal control at any time k does not only depend on the current state X_k^* but also on its the initial state ξ (via its mean).

Example: Mean-variance portfolio selection

The mean-variance discrete-time problem consists in minimizing the cost functional:

$$\begin{aligned} J(\alpha) &= \frac{\gamma}{2}\mathrm{Var}(X_n^\alpha) - \mathbb{E}[X_n^\alpha] \\ &= \mathbb{E}\Big[\frac{\gamma}{2}\big(X_n^\alpha\big)^2 - X_n^\alpha\Big] \; - \; \frac{\gamma}{2}\Big(\mathbb{E}[X_n^\alpha]\Big)^2, \end{aligned}$$

for some $\gamma > 0$, with a dynamics for the wealth process (X_k^α) valued in $E = \mathbb{R}$ controlled by the amount α_k valued in $A = \mathbb{R}$ invested in the stock at time k (we assume zero interest rate):

$$X_{k+1}^\alpha \;=\; X_k^\alpha + \alpha_k(b\Delta + \sigma\sqrt{\Delta}\varepsilon_{k+1}), \quad k = 0,\ldots,n-1, \; X_0^\alpha = x_0. \tag{4.17}$$

Here $x_0 \in \mathbb{R}$ is the initial capital, b, $\sigma > 0$ are some constants, representing respectively the rate of return and volatility of the stock, $\Delta > 0$ is a parameter, e.g. $\Delta = T/n$, arising when considering a time discretization of a continuous-time model over $[0,T]$, and (ε_k) is a sequence of i.i.d. random variables with distribution $\mathcal{N}(0,1)$. This univariate model fits into the LQ framework (4.4)-(4.5) with $d = m = 1$, and

$$\begin{aligned} &B_k = 1, \; \bar{B}_k = 0, \; C_k = b\Delta, \; \bar{C}_k = 0, \; D_k = \bar{D}_k = 0, \; H_k = \sigma\sqrt{\Delta}, \; \bar{H}_k = 0, \\ &Q_k = \bar{Q}_k = L_k = \bar{L}_k = R_k = \bar{R}_k = 0, \; Q = \frac{\gamma}{2}, \; \bar{Q} = -\frac{\gamma}{2}, \; L = 0, \; \bar{L} = -1. \end{aligned}$$

Conditions **(c0)-(c1)-(c2)** are clearly satisfied, and the Riccati system (4.14) for $(\Lambda_k, \Gamma_k, \rho_k, \chi_k)$ $\in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}$ is written in this case as:

$$\left\{ \begin{aligned} \Lambda_k &= \Lambda_{k+1}\tfrac{\sigma^2}{\sigma^2 + b^2\Delta} \\ \Gamma_k &= \tfrac{\sigma^2\Lambda_{k+1}}{\sigma^2\Lambda_{k+1} + b^2\Delta\Gamma_{k+1}}\Gamma_{k+1} \\ \rho_k &= \tfrac{\sigma^2\Lambda_{k+1}}{b^2\Delta\Gamma_{k+1} + \sigma^2\Lambda_{k+1}}\rho_{k+1} \\ \chi_k &= \chi_{k+1} - \tfrac{1}{4}\tfrac{b^2\Delta\rho_{k+1}^2}{\sigma^2\Lambda_{k+1} + b^2\Delta\Gamma_{k+1}}, \end{aligned} \right.$$

together with the terminal condition $\Lambda_n = \frac{\gamma}{2}$, $\Gamma_n = 0$, $\rho_n = -1$, $\chi_n = 0$. This leads by induction to the explicit form for $(\Lambda_k, \Gamma_k, \rho_k, \chi_k)$:

$$\left\{ \begin{aligned} \Lambda_k &= \tfrac{\gamma}{2}\Big(\tfrac{\sigma^2}{\sigma^2 + b^2\Delta}\Big)^{n-k}, \\ \Gamma_k &= 0, \;\; \rho_k \; = \; -1 \\ \chi_k &= -\tfrac{1}{2\gamma}\Big(\Big(\tfrac{\sigma^2 + b^2\Delta}{\sigma^2}\Big)^{n-k} - 1\Big). \end{aligned} \right. \tag{4.18}$$

The value functions are then explicitly given by

$$v_k(\mu) \;=\; \frac{\gamma}{2}\Big(\frac{\sigma^2}{\sigma^2 + b^2\Delta}\Big)^{n-k}\mathrm{Var}(\mu) - \bar{\mu} - \frac{1}{2\gamma}\Big(\Big(\frac{\sigma^2 + b^2\Delta}{\sigma^2}\Big)^{n-k} - 1\Big),$$

for all $k = 0,\ldots,n$, $\mu \in \mathcal{P}_2(\mathbb{R})$. Moreover, the optimal control is given in feedback form from (4.15) by:

$$\alpha_k^* \;=\; \tilde{\alpha}_k(X_k^*, \mathbb{P}_{X_k^*}) \;=\; -\frac{b}{\sigma^2 + b^2\Delta}\big(X_k^* - \mathbb{E}[X_k^*]\big) \;+\; \frac{b}{\sigma^2\gamma}\Big(\frac{\sigma^2 + b^2\Delta}{\sigma^2}\Big)^{n-k-1},$$

where $X_k^* = X_k^{\alpha^*}$ is the optimal wealth process with the feedback control α^*. It is then explicitly written from (4.16) by

$$\alpha_k^* \quad = \quad -\frac{b}{\sigma^2 + b^2\Delta}\Big[X_k^* - x_0 - \frac{1}{\gamma}\Big(1 + \frac{b^2}{\sigma^2}\Delta\Big)^n\Big]. \tag{4.19}$$

We then observe that the optimal control at any time k does not only depend on the current wealth X_k^* but also on the initial wealth x_0. This expression (4.19) of the optimal control is the discrete time analog of the continuous time optimal control obtained in [13] or [2]. Actually, if we view (4.17) as a time discretization (with a time step $\Delta = T/n$) of a continuous time Black-Scholes model for the stock price over $[0,T]$, with a controlled wealth dynamics

$$dX_t^\alpha \quad = \quad \alpha_t(bdt + \sigma dW_t), \quad X_0^\alpha = x_0,$$

then by sending n to infinity (hence Δ to zero) into (4.19), we retrieve the closed-form expression of the optimal control in [13] or [2]:

$$\alpha_t^* \quad = \quad -\frac{b}{\sigma^2}\Big[X_t^{\alpha^*} - x_0 - \frac{1}{\gamma}\exp\big(\frac{b^2}{\sigma^2}T\big)\Big].$$

References

[1] Ahmed N.U. and X. Ding (2001): "Controlled McKean-Vlasov equation", *Communications on Applied Analysis*, **5**, 183-206.

[2] Andersson D. and B. Djehiche (2010): "A maximum principle for SDEs of mean-field type", *Applied Mathematics and Optimization*, **63**, 341-356.

[3] Bensoussan A., Frehse J. and P. Yam (2013): *Mean Field Games and Mean Field Type Control Theory*, Springer Briefs in Mathematics.

[4] Bensoussan A., Frehse J. and P. Yam (2015): "The Master equation in mean-field theory", *Journal de Mathématiques Pures et Appliquées*, **103**(6), 1441-1474.

[5] Björk T. and A. Murgoci (2008): "A general theory of Markovian time inconsistent stochastic control problems", Preprint.

[6] Buckdahn R., Djehiche B. and J Li (2011): "A general maximum principle for SDEs of mean-field type", *Applied Mathematics and Optimization*, **64**(2), 197-216.

[7] Carmona R. and F. Delarue (2014): "Forward-backward Stochastic Differential Equations and Controlled McKean Vlasov Dynamics", to appear in the *Annals of Probability*.

[8] Carmona R., Delarue F. and A. Lachapelle (2013): "Control of McKean-Vlasov dynamics versus mean field games", *Mathematics and Financial Economics*, **7**, 131-166.

[9] Elliott R., Li X. and Y-H. Ni (2013): "Discrete time mean-field stochastic linear-quadratic optimal control problems", *Automatica*, **49**, 3222-3233.

[10] Huang M., Caines P., and R. Malhamé (2006): "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle", *Communication in Information and Systems*, **6**(3), 221-252.

[11] Lasry J.M. and P.L. Lions (2007): "Mean-field games", *Japanese Journal of Mathematics*, **2**, 229-260.

[12] Laurière M. and O. Pironneau (2014): "Dynamic programming for mean-field type control", CRAS, **352**(9), 707-713.

[13] Li D. and X.Y. Zhou (2000): "Continuous-time mean-variance portfolio selection: a stochastic LQ framework", *Applied Mathematics and Optimization*, **42**, 19-33.

[14] Pham H. and X.Wei (2015): "Bellman equation for McKean-Vlasov control problem", work in preparation.

[15] Sznitman A.S. (1989): Topics in propagation of chaos, in *Lecture Notes in Mathematics*, Springer, 1464, 165-251.

[16] Yong J. (2013): "A linear-quadratic optimal control problem for mean-field stochastic differential equations", *SIAM Journal on Control and Optimization*, **51(4)**, 2809-2838.